

## ΜΕΡΙΚΕΣ ΕΙΣΑΓΩΓΙΚΕΣ ΣΚΕΨΕΙΣ

### Φόβος και γοητεία

Η Τεχνητή Νοημοσύνη ήδη ασκούσε γοητεία, σε αδαείς και ειδικούς, πολύ πριν καταστεί αντικείμενο ευρύτατης δημόσιας συζήτησης. Ταυτόχρονα, συνδεδεμένη με την προσδοκία να εκτοξεύσει τις δυνατότητες της ανθρώπινης νοημοσύνης και να διευρύνει τα όρια κατανόησης (και τιθάσευσης;) του κόσμου, η ιδέα της υποκατάστασης «από τα ρομπότ» προκαλούσε, και εξακολουθεί να προκαλεί, δέος, αν όχι φόβο. Αναλογιζόμενος κανείς τις σκέψεις αλλά και τα συναισθήματα που γεννά η Sophia the Robot<sup>1</sup> όταν απευθύνεται στο Συμβούλιο των Ηνωμένων Εθνών, το chatbot της Google, που κλείνει ραντεβού για το κομμωτήριο<sup>2</sup>, ή τα συστήματα κοινωνικής βαθμολόγησης (social scoring) πολιτών στην Κίνα<sup>3</sup>, κατανοεί πως η γοητεία πάει χέρι χέρι με τον τρόμο<sup>4</sup>.

---

<sup>1</sup> J. Parviainen & M. Coeckelbergh, «The political choreography of the Sophia robot: beyond robot rights and citizenship to political performances for the social robotics market», *AI & Society* 36 (2021) 715-724.

<sup>2</sup> Για τη «συνομιλητική/διαλογική Τεχνητή Νοημοσύνη» βλ. K. Boonstra, «Introduction to Conversational AI», *The Definitive Guide to Conversational AI with Dialogflow and Google Cloud*, Μπέροκλεϋ: Apress, 2021.

<sup>3</sup> Η λεγόμενη «κοινωνική βαθμολόγηση» (social scoring) με χρήση συστημάτων Τεχνητής Νοημοσύνης εντάσσεται στις πρακτικές που η Ευρωπαϊκή Ένωση προτείνει να απαγορευθούν (βλ. άρθρο 5 της Πράξης για την Τεχνητή Νοημοσύνη). Για μια ανάλυση του κινεζικού συστήματος κοινωνικής βαθμολόγησης βλ. «The complicated truth about China's social credit system» σε <https://bit.ly/3l9JK4t>.

<sup>4</sup> Αντιστρέφοντας την παρατήρηση της Αγγέλας Καστρινακή στο *Από τον Φράνκενσταϊν στην Κλάρα: Παρατηρήσεις για την Τεχνητή Νοημοσύνη στη*

Πράγματι, αυτός ο «παροξυσμός αισιοδοξίας»<sup>5</sup> για την Τεχνητή Νοημοσύνη (εφεξής και ΤΝ) συνοδεύεται από —συχνά αντίστοιχα έντονους— προβληματισμούς για τους όρους, τα όρια και τις επιπτώσεις της ανάθεσης ανθρωπίνων δραστηριοτήτων σε τέτοια συστήματα. Η Siri και η Alexa είναι άραγε πολύτιμοι βοηθοί ή, παρόλο που μας την υποστήριξή τους, παίρνουν —μοιραία— και τον έλεγχο των πληροφοριών μας, άρα και της ζωής μας<sup>6</sup>; Το αλγοριθμικό VAR που θα αποφασίζει για την αξιολόγηση του οφσάιντ —και προοπτικά και των φάουλ— στους ποδοσφαιρικούς αγώνες θα μας στερήσει, εντέλει, πλήρως τις συχνά αμφιλεγόμενες αποφάσεις ή και την παρουσία του διαιτητή; Θα μπορούσε να εμπιστευτεί κανείς ανάλογα συστήματα στα δικαστήρια, αναφορικά με «απλές υποθέσεις», που θα υπόσχονταν επιτάχυνση της απονομής δικαιοσύνης<sup>7</sup>; Ένα σύστημα αξιολόγησης προσώπων (εργαζομένων, υποψήφιων δανειοληπτών κ.ά.) που βασίζεται σε αλγοριθμικό profiling είναι περισσότερο αξιόπιστο και αντικειμενικό από τον ανθρώπινο κριτή<sup>8</sup>; Εμπιστευόμαστε τον αλγόριθμο στην

---

*σύγχρονη τέχνη*, κείμενο που αποδίδει την εισήγηση σε εκδήλωση με τίτλο *Η Τεχνητή Νοημοσύνη και η φιλοσοφία της: Αναλύοντας το παρόν, ανιχνεύοντας το μέλλον*, η οποία διοργανώθηκε στις 25/5/2021 από τη Φιλοσοφική Σχολή του Πανεπιστημίου Κρήτης και δημοσιεύθηκε στο *The Books' Journal* 120 (2021) 16-18.

- 5 I. Σηφάκης, *Κατανοώντας και αλλάζοντας τον κόσμο*, Αθήνα: Αρμός, 2020, σ. 121.
- 6 Βλ. European Data Protection Board (Ευρωπαϊκό Συμβούλιο Προστασίας Δεδομένων — εφεξής και ΕΣΠΔ), Κατευθυντήριες γραμμές 2/2021 για τους εικονικούς φωνητικούς βοηθούς.
- 7 Βλ. ενδεικτικά N. Wang & M.Y. Tian, «“Intelligent Justice”: AI Implementations in China’s Legal Systems», στο A. Hanemaayer (επιμ.), *Artificial Intelligence and Its Discontents* («Social and Cultural Studies of Robots and AI»), Palgrave Macmillan, Cham, 2022. Επίσης, C. Shi, T. Sourdin & B. Li, «The Smart Court — A New Pathway to Justice in China?», *International Journal for Court Administration* (υπό δημοσίευση, 2021), προσβάσιμο στο SSRN: <https://bit.ly/3wffsZB>.
- 8 P. T. Kim & M. T. Bodie, «Artificial Intelligence and the Challenges of Workplace Discrimination and Privacy», *Journal of Labor and Employment Law* 35/2 (2021), 289-315.

αξιολόγηση των αιτήσεων για τα κοινωνικά επιδόματα ή μας ανησυχεί η «έλλειψη ενσυναίσθησης»<sup>9</sup>; Είμαστε σε θέση να κατανοήσουμε και να ελέγξουμε την Τεχνητή Νοημοσύνη ή μας αποθαρρύνει ήδη το ενδεχόμενο της (αναπόδραστης;) αυτονομήσής της; Τόσο ο ενθουσιασμός όσο και οι επιφυλάξεις τροφοδοτούνται από την ασάφεια που επικρατεί ως προς το τι ακριβώς είναι η Τεχνητή Νοημοσύνη.

Η Τεχνητή Νοημοσύνη φαίνεται πως δεν είναι απλώς «τεχνολογία γενικού σκοπού» που, όπως ο ατμός ή ο ηλεκτρισμός, μπορεί να χρησιμοποιηθεί σε απεριόριστο αριθμό δραστηριοτήτων· αλλά, πολύ περισσότερο, είναι τεχνολογία *gatekeeper*· μια τεχνολογία δηλαδή που, ως θυρωρός, φαίνεται να κατέχει τα κλειδιά για την εκθετική ανάπτυξη της ανθρώπινης ευημερίας. Αν και ανατρέχει στα μέσα της δεκαετίας του '50<sup>10</sup>, η αύξηση της υπολογιστικής ισχύος, η διαθεσιμότητα της πληροφορίας, η ανάλυση των μεγάλων δεδομένων και η πρόοδος —ως προς τους αλγορίθμους— εξελίσσουν πράγματι την ΤΝ στην πιο σημαντική τεχνολογία του 21ου αιώνα.

Αν και η σχετική συζήτηση είναι ιδιαίτερα εκτενής και πολλοί πλέον (θεωρούν ότι) έχουν κατανοήσει τι σημαίνει Τεχνητή Νοημοσύνη, δεν είναι διόλου ανώφελο να επιχειρήσουμε, αν όχι να την ορίσουμε, τουλάχιστον να αντιληφθούμε και να αποσαφηνίσουμε τι εννοούμε μιλώντας για Τεχνητή Νοημοσύνη. Για να είμαστε

<sup>9</sup> Αξίζει να σημειώσουμε ότι η καινοτομία της ΤΝ έχει πλέον επεκταθεί και στην αναζήτηση και εφαρμογή της ΤΝ με συναισθήματα και ενσυναίσθηση. Βλ. S. Marcos-Pablos & F.J. García-Peñalvo, «Emotional Intelligence in Robotics: A Scoping Review», στο *International Conference on Disruptive Technologies, Tech Ethics and Artificial Intelligence*, Springer Cham, 2021, σ. 66-75.

<sup>10</sup> Στην επιστημονική κοινότητα η Τεχνητή Νοημοσύνη εμφανίστηκε ως όρος για πρώτη φορά τη δεκαετία του 1950, όταν ο μαθηματικός Alan Turing δημοσίευσε το περίφημο άρθρο του «Computing Machinery and Intelligence». Λίγα χρόνια αργότερα, το 1956, ο John McCarthy έπλασε τον όρο «Τεχνητή Νοημοσύνη». Βλ. C. Smith, *The History of Artificial Intelligence*, University of Washington, 2006.

σαφείς, πρέπει να διευκρινίσουμε ότι η προσέγγισή μας να μεν —προφανώς— έχει ως αφετηρία τα τεχνολογικά δεδομένα, ωστόσο δεν επικεντρώνεται στην τεχνική διάσταση. Καθώς επιχειρούμε τη διάγνωση και ανάλυση ορισμένων πτυχών και επιπτώσεων της ΤΝ στα δικαιώματα, το δίκαιο και την εφαρμογή του, εκκινούμε από την κατανόηση που μπορεί να έχει για την ΤΝ όποιος επιφορτίζεται με την ερμηνεία, εφαρμογή αλλά και παραγωγή κανόνων και απαιτήσεων.

### *Τι είναι τελικά (ή αρχικά...) η Τεχνητή Νοημοσύνη;*

Στην κοινή αντίληψη, που έχει τραφεί εν πολλοίς και από τη λογοτεχνία και την επιστημονική φαντασία, η ΤΝ είναι συνώνυμο των ρομπότ<sup>11</sup>, που δεν είναι όμως απλώς «υπηρετές» του ανθρώπινου είδους, αλλά έχουν διανοητικές δυνατότητες οι οποίες υπερβαίνουν κατά πολύ εκείνες των ανθρώπων. Η εν λόγω πρόσληψη της ΤΝ συχνά διογκώνει τόσο τις θετικές όσο και τις αρνητικές διαστάσεις της και, περαιτέρω, φαίνεται να εμποδίζει την αντίληψή της ως παρούσα και όχι ως φουτουριστική πραγματικότητα.

Παρούσα είναι, ωστόσο, η Περιορισμένη Τεχνητή Νοημοσύνη (Narrow Artificial Intelligence)<sup>12</sup>, ενώ άδηλη είναι (ακόμη!) η χρονική στιγμή της μετάβασης στη λεγόμενη Γενική Τεχνητή Νοημοσύνη (General Artificial Intelligence)<sup>13</sup> και, ακόμη

<sup>11</sup> Δεν είναι λίγοι εκείνοι οι οποίοι την αντιμετωπίζουν ως καλύτερη και ταχύτερη υπολογιστική, ενώ άλλοι την ταυτίζουν με «ανθρωποειδή», χαρακτηριστικό που είναι ακριβές μόνο εν μέρει, καθώς η ρομποτική είναι μία έκφραση, ένα πεδίο της ΤΝ.

<sup>12</sup> Η περιορισμένη ή στενή Τεχνητή Νοημοσύνη μπορεί να επιλύσει πολύπλοκα προβλήματα πολύ γρήγορα. Ωστόσο, δεν διαθέτει αντίληψη για οτιδήποτε άλλο εκτός από τα δεδομένα που της παρέχουν οι δημιουργοί της και στο πλαίσιο της προκαθορισμένης από αυτούς λειτουργίας. Πρόκειται, δηλαδή, για ένα σύστημα που επιδεικνύει συγκεκριμένο επίπεδο νοημοσύνης στο πλαίσιο ενός αυστηρά καθορισμένου τομέα.

<sup>13</sup> Η Γενική Τεχνητή Νοημοσύνη αντιπροσωπεύει την Τεχνητή Νοημοσύνη «ανθρώπινου επιπέδου», με υπολογιστές εξίσου έξυπνους με τους ανθρώπους,

περισσότερο, στη λεγόμενη Τεχνητή Υπερνοημοσύνη (Artificial Superintelligence)<sup>14</sup>. Βέβαια, τα συστήματα ΤΝ ήδη ξεπερνούν τις ανθρώπινες δυνατότητες, καθώς μπορούν να υπολογίζουν ταχύτερα και ακριβέστερα και, υπό αυτήν την έννοια, συναγωνίζονται —συχνά με συντριπτικό πλεονέκτημα— τον άνθρωπο στην επίλυση προβλημάτων που απαιτούν διερεύνηση υπερβολικά μεγάλου αριθμού λύσεων ή συνδυασμό μεγάλου αριθμού προκαθορισμένων γνώσεων<sup>15</sup>. Χωρίς αμφιβολία, η ΤΝ αναγγέλλεται και εν μέρει πραγματώνεται ως η τεχνολογία που αντιλαμβάνεται ολιστικά το περιβάλλον και προσαρμόζει τις ενέργειες κατά τέτοιο τρόπο ώστε να μεγιστοποιεί τις πιθανότητες επιτυχίας<sup>16</sup>. Πρόκειται, επομένως, για το «εργαλείο» που προοπτικά θα υποκαταστήσει τις ανθρώπινες δραστηριότητες<sup>17</sup>.

Η ΤΝ συνιστά μια ολιστική, παγκόσμια και ταυτόχρονα αφηρημένη έννοια, που συμπεριλαμβάνει πολλούς και διαφορετικούς τρόπους λειτουργίας και πολλές και διαφορετικές εκδοχές, ο πυρήνας των οποίων δεν μπορεί ευχερώς να οριστεί ακόμη

---

σε κάθε πτυχή, και ικανούς να εκτελούν όλες τις πνευματικές διεργασίες τους.

- <sup>14</sup> Η Artificial Superintelligence δεν υπάρχει ακόμη. Αναμένεται ότι, μόλις καθιερωθεί η Γενική ΤΝ, θα εξελιχθεί σε Υπερνοημοσύνη πολύ γρήγορα, ως αποτέλεσμα μιας εκθετικής ανάπτυξης (φαινόμενο αποδιδόμενο ως «έκρηξη νοημοσύνης» ή «μοναδικότητα» ή ευρέως γνωστό με την αγγλική λέξη «singularity»).
- <sup>15</sup> Βλ. Ι. Σηφάκης, *Κατανοώντας και αλλάζοντας τον κόσμο*, ό.π., ο οποίος επισημαίνει ότι η ιδέα αυτή βρίσκει εφαρμογή στον Deep Blue και τον Watson της IBM, και τον AlphaGo της Google.
- <sup>16</sup> M. Janssen, M. Hartog, R. Matheus, A. Yi Ding & G. Kuk, «Will Algorithms Blind People? The Effect of Explainable AI and Decision-Makers' Experience on AI-supported Decision-Making in Government», *Social Science Computer Review* 40/2 (2021) 478-493, <https://bit.ly/3HaILGn> και περαιτέρω αναφορές σε S.J. Russell & P. Norvig, *Artificial intelligence: A modern approach*, Pearson, 2016.
- <sup>17</sup> S. Makridakis, «The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms», *Futures* 90 (Supplement C) (2017) 46-60.

επιστημονικά<sup>18</sup>. Δεν είναι όμως μια σαφώς ορισμένη τεχνολογία. Χωρίς αμφιβολία δε, εάν θέλαμε να ακριβολογούμε, δεν θα έπρεπε να μιλάμε για την Τεχνητή Νοημοσύνη αλλά για «πολλές τεχνητές νοημοσύνες»<sup>19</sup>, ενώ το περιεχόμενο της έννοιας αλλάζει με βάση το αν θα την αντιμετωπίσει κανείς ως επιστήμη<sup>20</sup> ή ως τεχνολογία. Μία άλλη δυσκολία ως προς τον ορισμό της συνίσταται ακριβώς στο ότι πρέπει κανείς να συμπεριλάβει την —πολυδιάστατη αλλά και απροσδιόριστη— έννοια της «ευφυΐας» για να προσδιορίσει τη «νοημοσύνη». Πρόκειται, σε κάθε περίπτωση, για μια έννοια δυναμική, με «στόχους» και «ορόσημα» διαρκώς μεταβαλλόμενα, οδηγώντας ορισμένους να ορίζουν τη «νοημοσύνη» ως «οτιδήποτε δεν έχουν κάνει ακόμη οι μηχανές», μια προσέγγιση που αναφέρεται ως το «θεώρημα του Tesler»<sup>21</sup>. Σε αυτήν ακριβώς τη σκέψη θεμελιώνεται και το λεγόμενο «AI effect»: Τεχνολογίες που χαρακτηρίζονταν ως ΤΝ παύουν να γίνονται αντιληπτές ως τέτοιες όταν χρησιμοποιούνται ευρέως και συνηθίζουμε στην εφαρμογή τους<sup>22</sup>. Ακόμη κι αν δεν ακολουθήσουμε αυτήν τη λογική, πρέπει να σημειώσουμε ότι, εντέλει, κάθε κατηγοριοποίηση είναι (εξ ορισμού;) προσωρινή, καθώς ευλόγως πιθανολογείται ότι

<sup>18</sup> M.R. Carrillo, «Artificial intelligence: From ethics to law», *Telecommunications Policy* 44/6 (2020) 101937.

<sup>19</sup> Οι διαφορετικοί τύποι, οι μεθοδολογίες και τα διαφορετικά συστήματα ΤΝ δεν εξυπηρετούν ενιαίους ή ίδιους σκοπούς και δεν πρέπει να αντιμετωπίζονται με τον ίδιο τρόπο. Βλ. Carrillo, ό.π.

<sup>20</sup> Ορισμένοι αναφέρονται σε αυτήν ως την ευρύτερη επιστήμη (αλλά και πρακτική) για την απόδοση ευφυΐας στις μηχανές. Βλ. B. W. Wirtz & W. M. Müller, «An integrated artificial intelligence framework for public management», *Public Manag. Rev.* 21/7 (2019), 1076-1100.

<sup>21</sup> R.J. Medaglia, R. Gil-Garcia & T. Pardo, «Artificial Intelligence in Government: Taking Stock and Moving Forward», *Social Science Computer Review* (July 2021) 1-18, doi: <https://bit.ly/3MBoPhd>.

<sup>22</sup> C. van Noordt & G. Misuraca, «Evaluating the impact of artificial intelligence technologies in public services: Towards an assessment framework», *Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance* (September 2022) 8-16, 12, doi: <https://bit.ly/3HaHP54>.

μπορεί να μην αποδίδει επαρκώς την τεχνολογική πραγματικότητα σε δέκα χρόνια από σήμερα.

Το κοινό σημείο των δεκάδων —ή και εκατοντάδων— ορισμών και προσεγγίσεων φαίνεται να εντοπίζεται στον συσχετισμό ή στην ταύτιση των συστημάτων ΤΝ με ανθρώπινες ή έλλογες ιδιότητες. Αν επιθυμούσαμε να υιοθετήσουμε μια απλή και περιεκτική προσέγγιση, θα μπορούσαμε να μιλήσουμε για την «τεχνολογία που επιτρέπει στους υπολογιστές να κάνουν πράγματα τα οποία απαιτούν νοημοσύνη όταν γίνονται από ανθρώπους», χωρίς να μας διαφεύγει ωστόσο ότι οι εφαρμογές και η στόχευση της ΤΝ υπερβαίνουν αυτόν τον λιτό ορισμό του Marvin Minsky<sup>23</sup>. Βέβαια, ένας τέτοιος ορισμός δεν ανταποκρίνεται στην απαίτηση για «ασφάλεια δικαίου», όταν τίθεται ζήτημα ρύθμισης και εφαρμογής (νομικών) κανόνων στις εφαρμογές ΤΝ.

Η Ευρωπαϊκή Ένωση, που αποδεικνύεται ιδιαίτερα ενεργή στον τομέα αυτόν, στην παρούσα φάση, φαίνεται να κατανοεί τόσο την απαίτηση για ασφάλεια και τη δημιουργία ενός «οικοσυστήματος εμπιστοσύνης» όσο και τις δυσκολίες και τα όρια του εγχειρήματος ήδη ως προς τον ορισμό: Η Ευρωπαϊκή Επιτροπή αποστασιοποιείται εν μέρει από την προσέγγιση του High-Level Expert Group on Artificial Intelligence (HLEG)<sup>24</sup> και επιλέγει έναν

<sup>23</sup> Για τον «κλασικό» ορισμό του Minsky βλ. C. Lexcellent, *Artificial Intelligence versus Human Intelligence*, Springer Briefs in Applied Sciences and Technology, σ. 5. Αντίστοιχος είναι ο ορισμός που προτείνει και η μελέτη του Συμβουλίου της Ευρώπης, η οποία αναφέρεται σε αλγοριθμικά μοντέλα που εκτελούν γνωστικές ή αντιληπτικές λειτουργίες, οι οποίες επιφυλάσσονταν σε σκεπτόμενα, κρίνοντα και έλλογα άτομα, βλ. D. Leslie, C. Burr, M. Aitken, J. Cows, M. Katell & M. Briggs, «Artificial intelligence, human rights, democracy, and the rule of law: A primer», 2021, διαθέσιμο σε SSRN 3817999.

<sup>24</sup> Το HLEG βασίζεται στην έννοια της ορθολογικότητας (rationality), που αναφέρεται στην ικανότητα να επιλέγεται η βέλτιστη ενέργεια προκειμένου να επιτευχθεί ένας στόχος, με βάση ορισμένα κριτήρια και τους διαθέσιμους πόρους. Σύμφωνα με τον εν λόγω ορισμό, πρόκειται για «τα συστήματα που επιδεικνύουν ευφυή συμπεριφορά μέσω της ανάλυσης του περιβάλλοντός τους και ανάληψης δράσης —με κάποιο βαθμό αυτονομίας—, για να

ορισμό που βασίζεται στα λειτουργικά χαρακτηριστικά του λογισμικού<sup>25</sup>. Στην Πρόταση κανονισμού για τη θέσπιση εναρμονισμένων κανόνων σχετικά με την Τεχνητή Νοημοσύνη (Πράξη για την Τεχνητή Νοημοσύνη), προτείνει να ορίζεται το «σύστημα ΤΝ» ως το «λογισμικό που έχει αναπτυχθεί με μία ή περισσότερες από τις τεχνικές και προσεγγίσεις που παρατίθενται στο Παράρτημα Ι<sup>26</sup> και μπορεί, για ένα δεδομένο σύνολο καθορισμένων από τον άνθρωπο στόχων, να παράγει αποτελέσματα όπως περιεχόμενο, προβλέψεις, συστάσεις ή αποφάσεις που επηρεάζουν τα περιβάλλοντα με τα οποία αυτές αλληλεπιδρούν». Κατά την πρόταση της Επιτροπής, ο κατάλογος αυτός θα πρέπει επικαιροποιείται από την ίδια διά «εκτελεστικών πράξεων», με βάση τις εξελίξεις της αγοράς και της τεχνολογίας<sup>27</sup>.

---

επιτύχουν συγκεκριμένους σκοπούς». Βλ. High Level Expert Group, A Definition of AI: Main Capabilities and Scientific Disciplines-Definition developed for the purpose of the deliverables of the High-Level Expert Group on AI Brussels, 18 December 2018.

- <sup>25</sup> «...[E]ιδικότερα στην ικανότητα ενός δεδομένου συνόλου στόχων καθορισμένων από τον άνθρωπο, να παράγει στοιχεία εξόδου όπως περιεχόμενο, προβλέψεις, συστάσεις ή αποφάσεις που επηρεάζουν το περιβάλλον με το οποίο αλληλεπιδρά το σύστημα, είτε σε υλική είτε σε ψηφιακή διάσταση...». Βλ. Σκέψη 6 του Προομίου της Πράξης για την Τεχνητή Νοημοσύνη (Πράξη για την Τεχνητή Νοημοσύνη – Artificial Intelligent ACT, εφεξής και ΑΙΑ).
- <sup>26</sup> Το Παράρτημα Ι συμπληρώνει τον ορισμό με τον κατάλογο προσεγγίσεων και τεχνικών για την ανάπτυξη της Τεχνητής Νοημοσύνης, που εμπεριέχουν: α) προσεγγίσεις μηχανικής μάθησης, όπως την εποπτευόμενη, μη εποπτευόμενη και ενισχυτική μάθηση, αξιοποιώντας μια ευρεία ποικιλία μεθόδων, συμπεριλαμβανομένης της βαθιάς μάθησης, β) προσεγγίσεις βασισμένες στη λογική και στη γνώση, περιλαμβάνοντας την αναπαράσταση της γνώσης, τον επαγωγικό (λογικό) προγραμματισμό, τις βάσεις γνώσεων, τα συμπεράσματα και τους αφαιρετικούς κινητήρες, (συμβολικά) συστήματα συλλογιστικής και εμπειρογνομώνων, και γ) στατιστικές προσεγγίσεις και μεθόδους εκτίμησης Bayesian, αναζήτησης και βελτιστοποίησης.
- <sup>27</sup> Σύμφωνα με το Σχέδιο Κανονισμού, η Επιτροπή εξουσιοδοτείται να εκδίδει «κατ' εξουσιοδότηση πράξεις» (delegated acts) για την ενημέρωση διαφόρων πτυχών του κανονισμού, προκειμένου να προσαρμόζονται στις εξελίξεις της αγοράς και της τεχνολογίας. Για παράδειγμα, εξουσιοδοτείται να τροποποιεί τον κατάλογο των προσεγγίσεων του Παραρτήματος Ι ή να επικαιροποιήσει



## Η «έλευση των μηχανών» και η αλλαγή (των κανόνων) του παιχνιδιού

Θέλουμε, εφόσον έχουμε τη δυνατότητα, να επιδιώξουμε υπολογιστές εξίσου νοήμονες με τον άνθρωπο, όπως πρότεινε το Turing test<sup>28</sup>; Ή περισσότερο νοήμονες; Ή διαφορετικά νοήμονες; Για να κάνουν τι ακριβώς και πώς; Και, εάν μπορούμε και θέλουμε, (πρέπει να) επιτρέπεται;

Διατυπώντας το αυτονόητο, ας σημειώσουμε κατ' αρχάς ότι δεν είναι η τεχνολογία καθεαυτή που αναπτύσσει επιπτώσεις, αλλά η χρήση της και ο τρόπος με τον οποίο επιδρά σε διαδικασίες και δομές της κοινωνίας, του κράτους, της οικονομίας. Συνεπώς, οι επιπτώσεις μιας τεχνολογίας διαφέρουν ανάλογα με τους σκοπούς που επιδιώκονται και το πλαίσιο εντός του οποίου αναπτύσσονται, χρησιμοποιούνται και επενεργούν<sup>29</sup>. Ισχύει αυτό και για την Τεχνητή Νοημοσύνη; Παρά το γεγονός ότι η TN συνδέεται με

---

τον κατάλογο των συστημάτων που προσδιορίζονται ως υψηλού κινδύνου. Η κανονιστική αρμοδιότητα της Επιτροπής συνιστά εν γένει πρόταση που αντιμετωπίζεται με δυσπιστία, καθώς, προκειμένου να διασφαλιστεί η ταχεία προσαρμογή στην τεχνολογική εξέλιξη, εμπλέκεται και στον ορισμό του πεδίου εφαρμογής (αναπόφευκτα) ένα όργανο (η Ευρωπαϊκή Επιτροπή) με κανονιστική εξουσιοδότηση (εκτελεστικές πράξεις), αλλά χωρίς τη δημοκρατική νομιμοποίηση των νομοθετικών οργάνων. Βέβαια, σύμφωνα με το άρθρο 73 παρ. 5 του Σχεδίου, μια κατ' εξουσιοδότηση πράξη ισχύει μόνο εάν το Ευρωπαϊκό Κοινοβούλιο ή το Συμβούλιο δεν διατυπώσει αντιρρήσεις εντός των προβλεπόμενων προθεσμιών.

- <sup>28</sup> Ο Alan Turing πρότεινε ένα τεστ (Turing test), για να κρίνουμε εάν ένας υπολογιστής είναι εξίσου νοήμων με έναν άνθρωπο. Το τεστ αυτό συνίσταται στη σύγκριση της συμπεριφοράς τους και των απαντήσεων που δίνουν στις ίδιες ερωτήσεις. Εάν ο ερωτών ερευνητής/κριτής δεν είναι σε θέση να ξεχωρίσει ποιος είναι ο άνθρωπος και ποιος ο υπολογιστής, τότε μπορούμε να θεωρήσουμε ότι είναι εξίσου νοήμονες.
- <sup>29</sup> Υποστηρίζεται ότι μερικοί τομείς της οικονομίας και του κράτους είναι πιο κατάλληλοι για την ανάπτυξη συστημάτων TN, όπως π.χ. η υγεία, η γεωργία ή οι μεταφορές. Βλ. W.G. de Sousa, E.R.P. de Melo, P.H.D.S. Bermejo, R.A.S. Farias & A.O. Gomes «How and where is artificial intelligence in the public sector going? A literature review and research agenda», *Government Information Quarterly* 36/4 (2019).

το φαινόμενο που αποκαλείται «έλευση των μηχανών» (rise of the machines), θα πρέπει να αναγνωρίσουμε ότι οι άνθρωποι εμπλέκονται σε κάθε στάδιο της ανάπτυξης και της εφαρμογής τεχνολογιών ΤΝ<sup>30</sup>.

Ωστόσο, τα συγκεκριμένα συστήματα, σύνθετα και όχι αμιγώς τεχνικά, χαρακτηρίζονται από πολυπλοκότητα και δυναμισμό: Τόσο οι υποκείμενοι μηχανισμοί όσο και η δυναμική και συνεχής αλληλεπίδραση με το περιβάλλον μέσα στο οποίο λειτουργούν είναι σύνθετα στη λειτουργική λογική τους, παράγοντας αποτελέσματα τα οποία είναι δύσκολο να προβλεφθούν, ιδίως όταν χρησιμοποιούνται αλγόριθμοι μηχανικής μάθησης. Η ΤΝ διαφέρει από τις κλασικές τεχνολογίες αυτοματισμού κατά το ότι δεν παράγει αποφάσεις επί τη βάση μιας προ-προγραμματισμένης λογικής «εάν-τότε» (if-then logic), όπου οι οδηγίες αναφορικά με το τι εισέρχεται παράγουν ακριβώς το ίδιο αποτέλεσμα. Αντίθετα, η ΤΝ χαρακτηρίζεται από αυτονομία στη λειτουργία και στην εκμάθηση<sup>31</sup>.

Η κατανόηση και πρόβλεψη της λειτουργίας τους στον πραγματικό κόσμο μπορεί να είναι ιδιαίτερα δυσχερής ακόμη και για όσους έχουν την απαραίτητη τεχνογνωσία. Καθίσταται δε ακόμη δυσχερέστερη λόγω της (εγγενούς) αδιαφάνειας, με αποτέλεσμα να μην είναι εφικτή η αξιολόγηση των κινδύνων —ακριβέστερα, ο εντοπισμός και η πρόληψη των επιβλαβών κοινωνικών επιπτώσεων και των διακυβευμάτων για τα δικαιώματα—, οι οποίοι επιτείνονται από τη διπλή φύση της υπολογιστικής αυτοματοποίησης, δηλαδή την αντικατάσταση της ανθρώπινης δραστηριότητας

<sup>30</sup> Για μια ενδιαφέρουσα αποτύπωση των χαρακτηριστικών της ΤΝ βλ. Κ. Yeung, *Responsibility and AI: A Study of the Implications of Advanced Digital Technologies (including AI Systems) for the Concept of Responsibility Within a Human Rights Framework; Summary and Conclusions*, Council of Europe, 2019.

<sup>31</sup> M.J. Ahn & Y.-C. Chen, «Artificial intelligence in government: Potentials, challenges, and the future», στο S.-J. Eom & J. Lee (επιμ.), *dg.o '20: The 21st annual international conference on digital government research*, Νέα Υόρκη: Association for Computing Machinery, 2020, σ. 243-252.

και την παραγωγή της νέας πληροφορίας. Η κλίμακα (του προβλήματος) είναι διαφορετική, καθώς η ΤΝ, σε συνδυασμό με την αναλυτική μεγάλων δεδομένων, συσσωρεύει, επεξεργάζεται και παράγει (νέα) πληροφορία με τρόπο και σε όγκο που οι άνθρωποι δεν μπορούν ούτε να φανταστούν. Το εν λόγω γεγονός ενδεχομένως δεν θα συνιστούσε πρόβλημα καθεαυτό, εάν δε συνοδευόταν από τη δυνατότητα χειραγώγησης των καταναλωτών μέσω της κατάρτισης προφίλ<sup>32</sup>, της άσκησης πολιτικής επιρροής σε εκατομμύρια ανθρώπους ή της διάδοσης αλγοριθμικών διακρίσεων σε κλίμακα χωρίς προηγούμενο<sup>33</sup>.

Η τεχνολογία είναι διαχρονικά μια πηγή αλλαγών, αλλά και αβεβαιοτήτων και κινδύνων<sup>34</sup>. Είναι αντιπροσωπευτικό πως ο χαρακτηρισμός «disruptive technology» —ένας όρος που μεταφράζεται ποικιλοτρόπως, ως «διασπαστικές τεχνολογίες» ή «ανatreπτικές τεχνολογίες», ανάλογα με τη στάση που ενυπάρχει στη θεώρησή μας<sup>35</sup>— αποδίδεται κατεξοχήν στην Τεχνητή Νοημοσύνη. Αν και ένας τέτοιος προσδιορισμός ενέχει αρνητική χροιά,

---

<sup>32</sup> P. Hacker, «Manipulation by Algorithms. Exploring the Triangle of Unfair Commercial Practice, Data Protection, and Privacy Law», *European Law Journal* (υπό έκδοση, 2021), σε <https://bit.ly/3vVD8ru>.

<sup>33</sup> Βλ. την πρόσφατη μελέτη του European Commission-European network of legal experts in gender equality and non-discrimination: J. Gerards & R. Xenidis, *Algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law*, EU Publications Office, 2021.

<sup>34</sup> Βλ. τη σχετική προβληματική ήδη στο κλασικό έργο του U. Beck, *Risikogesellschaft-Auf dem Weg in eine andere Moderne*, Suhrkamp Verlag 1986, καθώς επίσης A. Giddens, «Risk and Responsibility», *Modern Law Review* 62/1 (1999) 1-10 και Π. Μαντζούφα, *Συνταγματική Προστασία των Δικαιωμάτων στην Κοινωνία της Διακινδύνευσης*, Αθήνα-Θεσσαλονίκη: Εκδόσεις Σάκκουλα, 2006.

<sup>35</sup> Η διασπαστική ή ανατρεπτική τεχνολογία είναι μία τεχνολογία/καινοτομία που διαφοροποιεί τις αγορές και την οικονομία κατά τρόπο απροσδόκητο, βλ. R. Girassa, *Artificial Intelligence as a Disruptive Technology-Economic Transformation and Government Regulation*, Palgrave-Macmillan, 2020.

το βέβαιο είναι ότι οι «αλγόριθμοι»<sup>36</sup> αντιμετωπίζονται πλέον ως *game changers*.

### *Ο νομοθέτης και το δίλημμα του Collingridge*

Έχει κανόνες το νέο παιχνίδι; Η συζήτηση για την Τεχνητή Νοημοσύνη έθεσε εκ νέου και μετ' επιτάσεως το ζήτημα του δεοντολογικού προβληματισμού και των δικαιικών ή/και ηθικών αρχών και απαιτήσεων ως προς τον σχεδιασμό και τη χρήση της<sup>37</sup> – μια συζήτηση που είχε ήδη αρχίσει στην ΕΕ μέσω της ανάδειξης της ανάγκης για Υπεύθυνη Έρευνα και Καινοτομία (Responsible Research and Innovation)<sup>38</sup>. Η Ομάδα Εμπειρογνομώνων Υψηλού Επιπέδου για την Τεχνητή Νοημοσύνη (HLEG) υπογράμμισε ότι η συζήτηση για τις ηθικές αρχές και δεσμεύσεις «μπορεί να μας βοηθήσει να κατανοήσουμε τον τρόπο με τον οποίο η ανάπτυξη, η εγκατάσταση και η χρήση της ΤΝ μπορεί να επηρεάζει τα θεμελιώδη δικαιώματα και τις υποκειμένες αξίες τους, ενώ μπορεί να συμβάλει στην παροχή πιο λεπτομερούς καθοδήγησης κατά την εξέταση του τι θα πρέπει να κάνουμε αντί του τι μπορούμε να κάνουμε (επί του παρόντος) με την τεχνολογία»<sup>39</sup>.

Ο νομοθέτης δεν παράγει τεχνολογία. Παραμένει, όμως, υπεύθυνος για την ένταξη των νέων τεχνολογιών στην κοινωνία και

<sup>36</sup> Κατ' αρχάς, αντιλαμβανόμαστε έναν αλγόριθμο ως μια υπολογιστική διαδικασία ή ένα σύνολο μαθηματικών κανόνων που εφαρμόζονται σε μια τεχνολογία, προκειμένου να εκτελέσει μια εργασία ή/και να επιλύσει ένα πρόβλημα. Υπό την έννοια αυτή, εκπληρώνουν ανθρώπινους σκοπούς με έναν εξειδικεύσιμο και αναπαραξιμο τρόπο. Veale M. (2019), *Governing Machine Learning that Matters*, (Doctoral thesis, University College London, 2019).

<sup>37</sup> A. Tsamados, N. Aggarwal, J. Cows, J. Morley, H. Roberts, M. Taddeo & L. Floridi, «The ethics of algorithms: Key problems and solutions», *AI & Society* (2021) 1-16.

<sup>38</sup> V. Dignum, *Responsible artificial intelligence: How to develop and use AI in a responsible way*, Springer Nature, 2019, 49.

<sup>39</sup> HLEG (2019), Κατευθυντήριες Γραμμές Δεοντολογίας για Αξιοπίστη Τεχνητή Νοημοσύνη, 13.

για τη διαχείριση της σύνθετης έντασης μεταξύ των οικονομικών και κοινωνικών πλεονεκτημάτων και των κινδύνων που συνδέονται με αυτές. Η τεχνολογία ή, ακριβέστερα, οι συνέπειές της επιβάλλουν στον νομοθέτη να λάβει μια σειρά από σύνθετες αποφάσεις: εάν και πότε να παρέμβει, με ποιο είδος κανονιστικής παρέμβασης, ποιοι είναι οι εμπλεκόμενοι, πόσο πρέπει να διαρκέσει η κανονιστική παρέμβαση.

Κατ' αρχάς, ένα προκαταρκτικό ερώτημα είναι εάν πρέπει να παρέμβει ο νομοθέτης. Η «αρχή της φειδούς» θα επέβαλλε να απέχει από τη ρύθμιση, εάν οι υφιστάμενοι κανόνες είναι σε θέση να ρυθμίσουν ικανοποιητικά τα ζητήματα και τις συγκρούσεις που ανακύπτουν από την ΤΝ. Το ζήτημα τίθεται, για παράδειγμα, αναφορικά με την προστασία των προσωπικών δεδομένων όταν πραγματοποιείται αλγοριθμική επεξεργασία: είναι ή κατέστη ανεπαρκής η προστασία που προσφέρει ο Γενικός Κανονισμός Προστασίας Δεδομένων (εφεξής και ΓΚΠΔ), ένα δεσμευτικό νομοθέτημα που τέθηκε σε εφαρμογή το 2018, ακριβώς για να ανταποκριθεί στις προκλήσεις της ραγδαίας εξέλιξης της τεχνολογίας<sup>40</sup>; Αποστερείται το ισχύον δίκαιο τη ρυθμιστική ικανότητα όταν αλλάζει η τεχνολογία<sup>41</sup>;

Ακόμη και εάν τα προβλήματα που θέτει είναι πολυάριθμα και σημαντικά, θα πρέπει να υπογραμμίσουμε ότι έως τώρα η ΤΝ δεν αναπτύσσεται σε νομικό κενό<sup>42</sup>. Βέβαια, τα νομικά συστήματα

---

<sup>40</sup> L. Mitrou, «Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) “Artificial Intelligence-Proof”?», 2019, σε <https://bit.ly/3Mu0ee0>.

<sup>41</sup> Βλ. και P. De Hert, «Data Protection's Future without Democratic Bright Line Rules: Co-existing with Technologies in Europe after Breyer», EDPL 1/2017.

<sup>42</sup> Όπως επισήμαινε η Ευρωπαϊκή Επιτροπή στη Λευκή Βίβλο για την Τεχνητή Νοημοσύνη (2020), σημαντικό μέρος της υφιστάμενης νομοθεσίας της ΕΕ για την ασφάλεια των προϊόντων και την ευθύνη, συμπεριλαμβανομένων ειδικών τομεακών κανόνων, το οποίο συμπληρώνεται περαιτέρω από την εθνική νομοθεσία, είναι συναφές και ενδέχεται να εφαρμοστεί σε ορισμένες αναδυόμενες εφαρμογές της ΤΝ. Σε ό,τι αφορά την προστασία των θεμελιωδών

έχουν αναπτυχθεί έχοντας κατά νου ανθρώπους δρώντες ή ανθρώπους που λαμβάνουν αποφάσεις: τον τραπεζικό υπάλληλο που θα αξιολογήσει τη φερεγγυότητα, τον εργοδότη που αξιολογεί υποψηφίους προς εργασία, τον δημόσιο υπάλληλο που, σε ορισμένες περιπτώσεις, πρέπει να ασκήσει διακριτική ευχέρεια ή να συγχωρήσει τη λανθασμένη αίτηση για να χορηγήσει ένα επίδομα, ή τον «εξαντλημένο εφοριακό ο οποίος στο τέλος της ημέρας υπολογίζει λάθος μια επιστροφή φόρου»<sup>43</sup>. Οι εφαρμογές ΤΝ, υποστηρίζοντας ή υποκαθιστώντας την ανθρώπινη δράση και απόφαση, κάνουν παρόμοια ή διαφορετικά λάθη, προκαλούν ζημιά<sup>44</sup>, επιτείνουν τους παλαιούς κινδύνους ή θέτουν νέους. Ο νομοθέτης καλείται να αντιμετωπίσει το πρόβλημα του εύρους εφαρμογής και της αποτελεσματικής επιβολής της ισχύουσας νομοθεσίας, αλλά και εκείνο του διαρκούς εμπλουτισμού και μετασχηματισμού των συστημάτων ΤΝ, που αντίστοιχα δύναται να παραγάγει νέα ζητήματα και κινδύνους.

Έχουμε όμως γνώση και επίγνωση των νέων προβλημάτων και, πολύ περισσότερο, έναν βαθμό βεβαιότητας για την ενδε-

---

δικαιωμάτων και των δικαιωμάτων των καταναλωτών, το νομοθετικό πλαίσιο της ΕΕ περιλαμβάνει νομοθετικά μέσα, όπως την Οδηγία για τη φυλετική ισότητα, την Οδηγία για την ίση μεταχείριση στην απασχόληση και την εργασία, την Οδηγία για την ίση μεταχείριση ανδρών και γυναικών όσον αφορά την απασχόληση και την πρόσβαση σε αγαθά και υπηρεσίες· ακόμη, μια σειρά κανόνων προστασίας των καταναλωτών καθώς και κανόνες για την προστασία των δεδομένων προσωπικού χαρακτήρα και της ιδιωτικότητας, κυρίως τον Γενικό Κανονισμό Προστασίας Δεδομένων και άλλη τομεακή νομοθεσία που καλύπτει την προστασία των δεδομένων, όπως την Οδηγία προστασίας των δεδομένων για σκοπούς επιβολής του νόμου (σ. 16-17).

43 S. Ranchordas, «Experimental Regulations for AI: Sandboxes for Morals and Mores», University of Groningen Faculty of Law Research Paper Series 7 (2021) 5.

44 Σύμφωνα με την Ευρωπαϊκή Επιτροπή (Λευκή Βίβλος), η ζημία αυτή μπορεί να είναι υλική (ασφάλεια και υγεία των προσώπων, συμπεριλαμβανομένης της απώλειας ζωής, ζημία στην περιουσία) και άυλη (απώλεια της ιδιωτικότητας, περιορισμοί του δικαιώματος της ελευθερίας της έκφρασης και της ανθρώπινης αξιοπρέπειας, διακρίσεις, π.χ. όσον αφορά την πρόσβαση στην απασχόληση).

δειγμένη λύση; Το δίκαιο πρέπει να ανταποκρίνεται στις (μετ)αλλαγές της πραγματικότητας, αλλά οι αποτελεσματικές παρεμβάσεις προϋποθέτουν ότι κάποιος έχει κατανοήσει επαρκώς τα ζητήματα. Ως γενική παραδοχή, αξίζει να σημειωθεί ότι η καινοτομία είναι ένας «κινούμενος στόχος», που δεν συνδυάζεται εύκολα με το –συνήθως– αντιδραστικό μοντέλο και πλαίσιο της νομοθετικής παρέμβασης<sup>45</sup>. Η (εκ των πραγμάτων;) καθυστερημένη και συχνά ασθμαίνουσα (αντ)απόκριση του δικαίου στην τεχνολογία μοιάζει σχεδόν νομοτελειακή παραδοχή, καθώς η τεχνολογία, σε κάθε περίπτωση, εξελίσσεται ταχύτερα από τη νομοπαραγωγική διαδικασία<sup>46</sup>.

Θα μπορούσε, όμως, να είναι διαφορετικά; Και ποια είναι, εντέλει, η «κατάλληλη στιγμή»; Το δίλημμα που έθεσε ο Collingridge πριν από 40 χρόνια<sup>47</sup> παραμένει επίκαιρο: στα αρχικά στάδια εξέλιξης μιας τεχνολογίας, η ρύθμιση είναι δύσκολη ελλείψει επαρκούς πληροφορίας και γνώσης για τις συνέπειές της, ενώ, όταν διαγιγνώσκονται οι ανεπιθύμητες συνέπειές της, έχει πλέον διεισδύσει τόσο πολύ στην καθημερινότητα των ανθρώπων και στην οικονομία, που κάθε απόπειρα ελέγχου και ρύθμισης είναι δραστική, στοιχίζει και, ως εκ τούτου, συναντά την αντίσταση των παραγόντων που την αναπτύσσουν ή και επενδύουν σε αυτήν, αλλά συχνά και των ίδιων των χρηστών.

Η μηχανική μάθηση (machine learning)<sup>48</sup> και η ανάλυση δεδομένων (data analytics) έχουν ενταχθεί πλέον στη ζωή μας και

45 L. Bennett Moses, «How to think about law, regulation and technology – Problems with “technology” as a regulatory target», *Law, Innovation & Technology* 5/1 (2013) 1-20.

46 Για το θέμα αυτό βλ. Α. Μήτρου, *Το Δίκαιο στην Κοινωνία της Πληροφορίας*, Αθήνα-Θεσσαλονίκη: Εκδόσεις Σάκκουλα, 2002.

47 D. Collingridge, *The social control of technology*, Palgrave Macmillan, 1981, σ. 11.

48 Η μηχανική μάθηση συνιστά μεν ένα υπο-πεδίο της ΤΝ, αλλά συχνά ταυτίζεται με αυτήν. Πρόκειται για υπολογιστικά προγράμματα που μπορούν να «μαθαίνουν» μέσω της πληροφορίας και εμπειρίας, με την οποία τροφοδοτούνται από το «περιβάλλον» τους, και είναι σχεδιασμένα ώστε να βελτιώ-

διαθέτουν το δυναμικό να αλλάξουν τον τρόπο με τον οποίο συναλλασσόμαστε και επικοινωνούμε σε μια σειρά από επίπεδα: από την αυτόματη μετάφραση του Google translator και τη χρήση των εικονικών βοηθών, έως τη διάδοση συστημάτων αναγνώρισης προσώπου<sup>49</sup>, από τον τομέα της υγείας, έως τη χρήση τους από οικονομικές υπηρεσίες για την ανίχνευση της απάτης και την εισαγωγή αυτοκινούμενων οχημάτων χωρίς οδηγό. Από τις αρχές του 21ου αιώνα, και ιδιαίτερα κατά την τελευταία πενταετία, δεν παρατηρούμε μόνο την εκτίναξη των προσφερόμενων επιτευγμάτων και τη διεξόδυσή τους στην καθημερινότητα, αλλά και αντίστοιχη επέκταση του δημόσιου λόγου και διαλόγου.

### *Νόμος ή/και ηθική; Ρύθμιση ή/και ethics washing;*

Σε αυτό το πλαίσιο, όσο προφανής είναι η κοινή διαπίστωση ότι «κάτι πρέπει να γίνει», τόσο δυσχερής είναι ο προσδιορισμός του τι ακριβώς πρέπει να γίνει και πώς<sup>50</sup>. Με επίγνωση των δυσκολιών, αλλά και ως αποτέλεσμα σκεπτικισμού σχετικά με τις συνέπειες

---

νουν με αυτόν τον τρόπο την απόδοσή τους. Τα συστήματα αυτά μπορούν, δηλαδή, να μεταβάλλουν ή να τροποποιούν τη «συμπεριφορά» τους και να προσαρμόζονται, ώστε να αυξάνουν την επίδοσή τους. Η μηχανική μάθηση διακρίνεται σε επιβλεπόμενη και μη επιβλεπόμενη. Στην τελευταία περίπτωση ο αλγόριθμος δεν γνωρίζει το ορθό αποτέλεσμα ή τον τρόπο ταξινόμησης των δεδομένων εισόδου, αλλά, με βάση ένα σύνολο μη επεξεργασμένων δεδομένων, εντοπίζει τα σημαντικά χαρακτηριστικά σε αυτά τα δεδομένα και μαθαίνει πώς να τα ταξινομεί στις κατάλληλες κατηγορίες.

- <sup>49</sup> Το σύστημα αναγνώρισης προσώπου (facial recognition systems) βρίσκονται στο επίκεντρο της δημόσιας συζήτησης και αντιπαράθεσης. Α.α. βλ. European Parliament, Regulating facial recognition in the EU, September 2021. Προσφάτως παρουσιάστηκε η εφαρμογή με βάση την οποία θα μπορεί να «απεικονιστεί», κατ' ουσίαν να «δημιουργηθεί», το πρόσωπο του υπόπτου ή του θύματος από τον συνδυασμό ανάλυσης γενετικού υλικού και χρήσης βάσεων δεδομένων με εικόνα.
- <sup>50</sup> B. Wagner, «Ethics as an escape from regulation: From “ethics-washing” to ethics-shopping?», στο E. Bayamlıoğlu, I. Baraliuc & L. Janssens (επιμ.), *Being Profiled: Cogitas Ergo Sum. 10 Years of Profiling the European Citizen*, Amsterdam University Press, 2018, σ. 84-88.



της ρυθμιστικής παρέμβασης στην καινοτομία και, εντέλει, στην ανάπτυξη, προτεινόταν (και εξακολουθεί να προτείνεται) η εστίαση στην καθιέρωση και αποδοχή μιας «ηθικής της Τεχνητής Νοημοσύνης»<sup>51</sup> ή ο προσανατολισμός στην «ηθική Τεχνητή Νοημοσύνη» —μια διατύπωση που αποδίδει μεν την ανθρωποκεντρική (ή και ανθρωπομορφική) πρόσληψη της Τεχνητής Νοημοσύνης, αλλά αγνοεί ότι πρόκειται για σύνθετα συστήματα που συναποτελούνται από τεχνολογίες, ανθρώπους και τις ανάγκες και στόχους αυτών<sup>52</sup>.

Η ηθική της Τεχνητής Νοημοσύνης συνιστά πρωτίστως μια ηθική της τεχνολογίας (technology ethics). Ο σχεδιασμός και η ανάπτυξη συστημάτων και εφαρμογών δύναται να απολήξει σε προσβολές δικαιωμάτων, σε συγκρούσεις αξιών και κανόνων· αλλά και αντιστρόφως, χωρίς αμφιβολία, οι αξιακές παράμετροι και οι αντίστοιχες δεσμεύσεις επηρεάζουν (ή μπορούν να επηρεάσουν) τα χαρακτηριστικά του τεχνικού σχεδιασμού. Ωστόσο, καθώς οι τεχνολογίες ΤΝ ενσωματώνονται ως εργαλεία για τη βελτιστοποίηση της αποτελεσματικότητας του κοινωνικού συντονισμού, (θα) παράγουν αποτελέσματα και αποφάσεις που δίνουν προτεραιότητα σε μερικές αξίες έναντι άλλων και (θα) επηρεάζουν άμεσα πρόσωπα και ομάδες. Βέβαια, η επίτευξη της σχεδίασης ενός τεχνολογικού προϊόντος που ενσωματώνει αρχές δεν απαιτεί μόνο σχεδιαστική δεινότητα, αλλά και αναστοχαστική κατανόηση των αντίστοιχων αξιών και επιλογών και του τρόπου με τον οποίο αυτές επιδρούν στη ζωή και τα δικαιώματα των ατόμων<sup>53</sup>. Τα συστήματα

---

<sup>51</sup> K. van Dijk, S. Casiraghi & S. Gutwirth, «The “Ethification” of ICT Governance: Artificial Intelligence and Data Protection in the European Union», *Computer Law and Security Review* 43 (2021).

<sup>52</sup> Βλ. την ενδιαφέρουσα ανάλυση της Dignum, ό.π.

<sup>53</sup> Αναφορικά με τη σχέση της τεχνολογίας με τις κοινωνικές, ηθικές και πολιτικές αξίες βλ. ενδεικτικά U. Pagallo, «On the Principle of Privacy by Design and its Limits: Technology, Ethics and the Rule of Law», στο S. Gutwirth, R. Leenes, P. De Hert & Y. Poullet (επιμ.), *European Data Protection: In Good Health?*, Springer, 2012, σ. 331-346, καθώς και M. Flanagan, M. Howe, C. Daniel & H. Nissenbaum, «Embodying values in technology: Theory and

TN ανταναικλούν τις αξίες και τις προτεραιότητες αυτών που τα αναπτύσσουν, αλλά ενδέχεται να μην ευθυγραμμίζονται με τις συλλογικές –δημοκρατικές και συνταγματικές– αξίες<sup>54</sup>.

Τόσο στην Ευρωπαϊκή Ένωση<sup>55</sup> όσο και σε διεθνές επίπεδο, διατυπώθηκαν ηθικές και δεοντολογικές αρχές στις οποίες καλούνται να δεσμευτούν όσοι δημιουργούν αλλά και όσοι χρησιμοποιούν συστήματα TN. Οι πρωτοβουλίες είχαν αφετηρία είτε ιδιωτικούς φορείς, που έτειναν να υποβαθμίζουν τον ρόλο του κράτους και να αναδεικνύουν αυτόν του ιδιωτικού τομέα, είτε διεθνείς οργανισμούς, που αποσκοπούσαν στην αποτύπωση ενός δικαιιοπολιτικού πλαισίου<sup>56</sup> με παγκόσμιους αποδέκτες παραγωγούς και παρόχους TN, μη υποκείμενους στα όρια μιας εθνικής επικράτειας και κανονιστικής δικαιοδοσίας. Οι αρχές που έχουν διατυπωθεί παρουσιάζουν σημεία τομής, αλλά διαφέρουν μεταξύ τους τόσο ως προς τη γεωγραφικά προσδιορισμένη εμβέλειά τους όσο και ως προς το περιεχόμενο<sup>57</sup>. Επίσης, ελλείψει δεσμευτικών ρυθμίσεων, ενδεχομένως να καταλήξουν να ενθαρρύνουν και το λεγόμενο *ethics shopping*, δηλαδή την πρακτική του συνδυασμού (*mix and match*) διάφορων ηθικών αρχών, κυρίως προκειμένου να αιτιολογηθούν εκ των υστέρων προ-υφιστάμενες πρακτικές και συμπεριφορές, αντί να προωθηθούν νέες επί τη βάση των ηθικών αρχών<sup>58</sup>.

---

practice», στο J. van den Hoven & J. Weckert (επιμ.), *Information technology and moral philosophy*, Cambridge University Press, 2008, σ. 322-353.

<sup>54</sup> Βλ. Yeung, ό.π.

<sup>55</sup> HLEG, Κατευθυντήριες Γραμμές Δεοντολογίας για Αξιόπιστη Τεχνητή Νοημοσύνη, 2019.

<sup>56</sup> Το 2019 ήταν μια ιδιαίτερα σημαντική χρονιά, καθώς το Συμβούλιο της Ευρώπης δημοσίευσε τις *Guidelines on Artificial Intelligence and Data Protection*, ενώ ο ΟΟΣΑ υιοθέτησε τις *OECD Recommendation on AI* και ακολούθησε αντίστοιχη δήλωση των G20. Το 2021 η UNESCO παρουσίασε τις *Recommendation on the Ethics of Artificial Intelligence*.

<sup>57</sup> R. Clarke, *The EC's Proposal for Regulation of AI: Evaluation against a Consolidated Set of 50 Principles*, 2021.

<sup>58</sup> European Group on Ethics in Science and New Technologies, «Statement on Artificial Intelligence, Robotics and “Autonomous” Systems»,

Η ανάδειξη των κανόνων ηθικής και δεοντολογίας —ως βασικού ρυθμιστικού εργαλείου— σχετίζεται και με την προσπάθεια να διατυπωθούν μεν οι δικαιοπολιτικοί στόχοι και οι συναφείς δεσμεύσεις, όπως η απαγόρευση των διακρίσεων<sup>59</sup> ή η διαφάνεια, αλλά να μη λειτουργήσουν ως εμπόδιο στην εξέλιξη των νέων τεχνολογιών<sup>60</sup>. Η επίκληση του «ηθικού πλαισίου» σχεδιασμού και χρήσης, που περιβάλλεται έναν οιονεί κανονιστικό μανδύα, παρέχει το επιχείρημα της υπέρβασης των υφιστάμενων κανονιστικών ορίων και τη δυνατότητα να τα αγνοήσει κανείς<sup>61</sup>. Ο αριθμός των ηθικών αρχών και των δεοντολογικών κατευθύνσεων/οδηγιών βαίνει διαρκώς αυξανόμενος<sup>62</sup>, επιβεβαιώνοντας την ενεργή κινητοποίηση για την πλαισίωση της ΤΝ με «κανόνες», αλλά και την επιδίωξη να αποφευχθεί —ή έστω να καθυστερήσει— η δεσμευτική κανονιστική παρέμβαση του νομοθέτη. Δεν είναι λίγοι εκείνοι που αναφέρονται σε αυτή την προσπάθεια να διατηρηθεί η απόπειρα ρύθμισης της ΤΝ στο επίπεδο της ηθικής δέσμευσης με τον χαρακτηρισμό «ethics wash(ing)»<sup>63</sup>.

---

Luxembourg, Publication Office of the European Union, 2018, 5. Επίσης σχετικά N. Van Dijk, S. Casiraghi & S. Gutwirth, «The “Ethification” of ICT Governance. Artificial Intelligence and Data Protection in the European Union», *Computer Law & Security Review*, 43 (2021).

- <sup>59</sup> Για το μείζον ζήτημα των αλγοριθμικών διακρίσεων βλ. τη μελέτη της Ηλ. Κωστή στον παρόντα νόμο. Επίσης, την πρόσφατη μελέτη του European Commission, European network of legal experts in gender equality and non-discrimination: J. Gerards & R. Xenidis, *Algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law*, 2021.
- <sup>60</sup> C. Klover & A. Fanta, «No red lines: Industry defuses ethics guidelines for artificial intelligence», 2019, σε <https://bit.ly/3QeaA4L>.
- <sup>61</sup> B. Wagner, «Ethics as an Escape from Regulation: From ethics-washing to ethics-shopping?», στο M. Hildebrandt (επιμ.), *Being Profiling. Cogitas ergo sum*, Amsterdam University Press, 2018.
- <sup>62</sup> C. Stix, «Actionable principles for artificial intelligence policy: Three pathways», *Science and engineering ethics*, 27/1 (2021) 1-17.
- <sup>63</sup> Επίσης, J. Morley, A. Elhalal, F. Garcia, L. Kinsey, J. Mökander, L. Floridi, «Ethics as a service: A pragmatic operationalisation of AI ethics», *Minds*

Ακόμη κι αν παρακάμψουμε τις αιτιάσεις για κατάχρηση της ηθικής, δεν μπορούμε να αποφύγουμε το εξής ερώτημα: αρκεί η δέσμευση, με βάση τα (ρευστά) Artificial Intelligence ethics<sup>64</sup>, σε ένα —εξ ορισμού αόριστο;— soft law, για να διασφαλιστεί ότι δεν προσβάλλονται θεμελιώδη δικαιώματα και αρχές; Στα διάφορα πλαίσια ρύθμισης, οι αρχές αυτές είναι συνήθως διατυπωμένες με γενικό και αφηρημένο τρόπο και είναι αμφίβολο, για παράδειγμα, εάν οι δημιουργοί τέτοιων συστημάτων δύνανται να συνάγουν ακριβέστερες οδηγίες ως προς το πώς θα σχεδιάζουν ένα σύστημα ΤΝ, ώστε να συμμορφώνονται με αυτές<sup>65</sup>. Κυρίως, όμως, δεν υπάρχουν αναφορές σε μηχανισμούς και όργανα που θα εξασφάλιζαν την εφαρμογή τους<sup>66</sup>. Η «ηθική ρύθμιση» ή η «ηθική δέσμευση» δεν διαθέτει από τη φύση της τα εργαλεία που θα την καθιστούσαν αποτελεσματική<sup>67</sup>.

Ένα άλλο μείζον ζήτημα αφορά στην αντιπροσωπευτικότητα των «ηθικών αξιών» στις οποίες καλούνται ή είναι διατεθειμένοι να δεσμευτούν οι δημιουργοί και πάροχοι. Όπως επισημαίνεται,

---

*and Machines* 31/2 (2021) 239-256. Σχετικά βλ. και το κείμενο του Σπ. Τάσση στον παρόντα τόμο.

<sup>64</sup> Η συζήτηση δεν είναι νέα. Ήδη, πριν από πολλές δεκαετίες, αντιτασσόταν στον ενστερνισμό κανόνων προστασίας δεδομένων η υιοθέτηση κανόνων δεοντολογίας και εν γένει soft law.

<sup>65</sup> R. Clarke, «Principles and business processes for responsible AI», *Computer Law & Security Review* 35/4 (2019) 410-422.

<sup>66</sup> Εξάιρεση αποτελεί, υπό μία έννοια, η Σύσταση του Συμβουλίου της Ευρώπης για τις επιπτώσεις των αλγοριθμικών συστημάτων στα ανθρώπινα δικαιώματα, που αναφέρεται σε πιο συγκεκριμένες υποχρεώσεις ως προς την πρόβλεψη μέτρων για τη διασφάλιση της εφαρμογής των δεοντολογικών αρχών. Βλ. Recommendation to member States on the human rights impacts of algorithmic systems, Council of Europe Committee of Ministers, CM/Rec (2020). Πρόκειται για ένα λεπτομερές κείμενο που απευθύνεται τόσο στα κράτη όσο και σε ιδιώτες.

<sup>67</sup> A. Ressaygues & R. Rodrigues, «AI ethics should not remain toothless! A call to bring back the teeth of ethics», *Big Data & Society* 7/2 (2020). Επίσης T. Hagendorff, «The ethics of AI ethics: An evaluation of guidelines», *Minds and Machines* 30 (2020) 99-120.

ακόμη και σε σχέση με συστήματα που μπορούν να θίγουν άμεσα και διεισδυτικά τους ανθρώπους<sup>68</sup>, δεν έχουν —είτε ατομικά είτε συλλογικά— ουσιαστικά την ευκαιρία στον προσδιορισμό των αξιών, των στόχων ή των ισορροπιών που τα εν λόγω συστήματα είναι σχεδιασμένα να εκφράσουν<sup>69</sup>. Οι πιθανές αρνητικές επιπτώσεις επιτείνονται, εξαιτίας της μεγάλης ασυμμετρίας ισχύος μεταξύ εκείνων που αναπτύσσουν και εφαρμόζουν αλγοριθμικά συστήματα και των χρηστών ή όσων υφίστανται τα αποτελέσματα αυτών των συστημάτων. Η ασυμμετρία αυτή δεν εντοπίζεται μόνο στις σχέσεις μεταξύ μεγάλων ιδιωτών «παικτών» που παράγουν τέτοια συστήματα και των κρατικών οντοτήτων που τα χρησιμοποιούν, ή μεταξύ παρόχων και ιδιωτών χρηστών<sup>70</sup>. Η παραδοσιακή ασυμμετρία πληροφορίας και ισχύος μεταξύ κράτους και πολιτών<sup>71</sup> εξελίσσεται σε ασυμμετρία μεταξύ όσων εφαρμόζουν αλγορίθμους και εκείνων που κυβερνώνται δι' αυτών<sup>72</sup>.

<sup>68</sup> Όπως συστήματα και πρακτικές που, κατά το Σχέδιο Πράξης για την Τεχνητή Νοημοσύνη, εντάσσονται στην κατηγορία μη αποδεκτού κινδύνου και απαγορεύονται. Τέτοιες πρακτικές είναι π.χ. η χρήση τεχνικών που απευθύνονται στο υποσυνείδητο, η χρήση συστημάτων ΤΝ από τις δημόσιες αρχές ή εξ ονόματός τους για την αξιολόγηση ή την ταξινόμηση της αξιοπιστίας των φυσικών προσώπων για ορισμένο χρονικό διάστημα, με βάση την κοινωνική τους συμπεριφορά ή τα γνωστά ή προβλεπόμενα προσωπικά χαρακτηριστικά τους ή χαρακτηριστικά της προσωπικότητάς τους, με κοινωνική βαθμολογία.

<sup>69</sup> Yeung, ό.π.

<sup>70</sup> Όπως επισημαίνεται στην έκθεση του Συμβουλίου της Ευρώπης, οι πάροχοι τέτοιων συστημάτων και υπηρεσιών ενισχύουν την υπεροχή τους έναντι των χρηστών μέσω της ικανότητας να προβαίνουν σε διεισδυτική επιτήρησή τους.

<sup>71</sup> Μία ασυμμετρία που είχε επισημάνει ο S. Simitis, ήδη τη δεκαετία του '60, και η οποία οδήγησε και στη συνειδητοποίηση της ανάγκης για υιοθέτηση κανόνων για την προστασία των προσωπικών δεδομένων. Βλ. S. Simitis, *Informationskrise des Rechts und Datenverarbeitung*, C.F. Müller Verlag, 1970.

<sup>72</sup> B. Wagner, «Study on The Human Rights Dimensions of Automated Data Processing Techniques (In Particular Algorithms) And Possible Regulatory Implications», 6 October, Council of Europe, Committee of Experts on internet intermediaries (MSI-NET), 2017. Σχετικά βλ. και το κείμενο της γράφουσας στον παρόντα τόμο.